● ●

```
    1 CCCGGGCGAG GACCCCTCCA GGATGCAGGT CCCGAACAGC ACCGGCCCGG
   51 ACAACGCGAC GCTGCAGATG CTGCGGAACC CGGCGATCGC GGTGGCCCTG
  101 CCCGTGGTGT ACTCGCTGGT GGCGGCGGTC AGCATCCCGG GCAACCTCTT
  151 CTCTCTGTGG GTGCTGTGCC GGCGCATGGG GCCCAGATCC CCGTCGGTCA
  201 TCTTCATGAT CAACCTGAGC GTCACGGACC TGATGCTGGC CAGCGTGTTG
  251 CCTTTCCAAA TCTACTACCA TTGCAACCGC CACCACTGGG TATTCGGGGT
  301 GCTGCTTTGC AACGTGGTGA CCGTGGCCTT TTACGCAAAC ATGTATTCCA
  351 GCATCCTCAC CATGACCTGT ATCAGCGTGG AGCGCTTCCT GGGGGGTCCTG
  401 TACCCGCTCA GCTCCAAGCG CTGGCGCCGC CGTCGTTACG CGGTGGCCGC
  451 GTGTGCAGGG ACCTGGCTGC TGCTCCTGAC CGCCCTGTCC CCGCTGGCGC
  501 GCACCGATCT CACCTACCCG GTGCACGCCC TGGGCATCAT CACCTGCTTC
  551 GACGTCCTCA AGTGGACGAT GCTCCCCAGC GTGGCCATGT GGGCCGTGTT
  601 CCTCTTCACC ATCTTCATCC TGCTGTTCCT CATCCCGTTC GTGATCACCG
  651 TGGCTTGTTA CACGGCCACC ATCCTCAAGC TGTTGCGCAC GGAGGAGGCG
  701 CACGGCCGGG AGCAGCGGAG GCGCGCGGTG GGCCTGGCCG CGGTGGTCTT
  751 GCTGGCCTTT GTCACCTGCT TCGCCCCCAA CAACTTCGTG CTCCTGGCGC
  801 ACATCGTGAG CCGCCTGTTC TACGGCAAGA GCTACTACCA CGTGTACAAG
  851 CTCACGCTGT GTCTCAGCTG CCTCAACAAC TGTCTGGACC CGTTTGTTTA
  901 TTACTTTGCG TCCCGGGAAT TCCAGCTGCG CCTGCGGGAA TATTTGGGCT
  951 GCCGCCGGGT GCCCAGAGAC ACCCTGGACA CGCGCCGCGA GAGCCTCTTC
 1001 TCCGCCAGGA CCACGTCCGT GCGCTCCGAG GCCGGTGCGC ACCCTGAAGG
 1051 GATGGAGGGA GCCACCAGGC CCGGCCTCCA GAGGCAGGAG AGTGTGTTCT
 1101 GAGTCCCGGG GGCGCAGCTT GGAGAGCCGG GGGCGCAGCT TGGAGATCC
 1151 GGGGCGCATG GAGAGGCCAC GGTGCCAGAG GTTCAGGGAG AA
```

FIG. 1

FIG. 2.   Predict d polypeptide sequence of HG52.

```
  1   MQVPNSTGPD  NATLQMLRNP  AIAVALPVVY  SLVAAVSIPG  NLFSLWVLCR

 51   RMGPRSPSVI  FMINLSVTDL  MLASVLPFQI  YYHCNRHHWV  FGVLLCNVVT

101   VAFYANMYSS  ILTMTCISVE  RFLGVLYPLS  SKRWRRRRYA  VAACAGTWLL

151   LLTALSPLAR  TDLTYPVHAL  GIITCFDVLK  WTMLPSVAMW  AVFLFTIFIL

201   LFLIPFVITV  ACYTATILKL  LRTEEAHGRE  QRRRAVGLAA  VVLLAFVTCF

251   APNNFVLLAH  IVSRLFYGKS  YYHVYKLTLC  LSCLNNCLDP  FVYYFASREF

301   QLRLREYLGC  RRVPRDTLDT  RRESLFSART  TSVRSEAGAH  PEGMEGATRP

351   GLQRQESVF
```

3/6

```
          10                    30                    50
CCGGGCGAGGACCCCTCCAGGATGCAGGTCCCGAACAGCACCGGCCCGGACAACGCGACG
                          MetGlnValProAsnSerThrGlyProAspAsnAlaThr


          70                    90                    110
CTGCAGATGCTGCGGAACCCGGCGATCGCGGTGGCCCTGCCCGTGGTGTACTCGCTGGTG
LeuGlnMetLeuArgAsnProAlaIleAlaValAlaLeuProValValTyrSerLeuVal


          130                   150                   170
GCGGCGGTCAGCATCCCGGGCAACCTCTTCTCTCTGTGGGTGCTGTGCCGGCGCATGGGG
AlaAlaValSerIleProGlyAsnLeuPheSerLeuTrpValLeuCysArgArgMetGly


          190                   210                   230
CCCAGATCCCCGTCGGTCATCTTCATGATCAACCTGAGCGTCACGGACCTGATGCTGGCC
ProArgSerProSerValIlePheMetIleAsnLeuSerValThrAspLeuMetLeuAla


          250                   270                   290
AGCGTGTTGCCTTTCCAAATCTACTACCATTGCAACCGCCACCACTGGGTATTCGGGGTG
SerValLeuProPheGlnIleTyrTyrHisCysAsnArgHisHisTrpValPheGlyVal


          310                   330                   350
CTGCTTTGCAACGTGGTGACCGTGGCCTTTTACGCAAACATGTATTCCAGCATCCTCACC
LeuLeuCysAsnValValThrValAlaPheTyrAlaAsnMetTyrSerSerIleLeuThr


          370                   390                   410
ATGACCTGTATCAGCGTGGAGCGCTTCCTGGGGGTCCTGTACCCGCTCAGCTCCAAGCGC
MetThrCysIleSerValGluArgPheLeuGlyValLeuTyrProLeuSerSerLysArg


          430                   450                   470
TGGCGCCGCCGTCGTTACGCGGTGGCCGCGTGTGCAGGGACCTGGCTGCTGCTCCTGACC
TrpArgArgArgArgTyrAlaValAlaAlaCysAlaGlyThrTrpLeuLeuLeuLeuThr


          490                   510                   530
GCCCTGTCCCCGCTGGCGCGCACCGATCTCACCTACCCGGTGCACGCCCTGGGCATCATC
AlaLeuSerProLeuAlaArgThrAspLeuThrTyrProValHisAlaLeuGlyIleIle


          550                   570                   590
ACCTGCTTCGACGTCCTCAAGTGGACGATGCTCCCCAGCGTGGCCATGTGGGCCGTGTTC
ThrCysPheAspValLeuLysTrpThrMetLeuProSerValAlaMetTrpAlaValPhe
```

Fig.3A

4/6

```
           610                 630                 650
CTCTTCACCATCTTCATCCTGCTGTTCCTCATCCCGTTCGTGATCACCGTGGCTTGTTAC
LeuPheThrIlePheIleLeuLeuPheLeuIleProPheValIleThrValAlaCysTyr

           670                 690                 710
ACGGCCACCATCCTCAAGCTGTTGCGCACGGAGGAGGCGCACGGCCGGGAGCAGCGGAGG
ThrAlaThrIleLeuLysLeuLeuArgThrGluGluAlaHisGlyArgGluGlnArgArg

           730                 750                 770
CGCGCGGTGGGCCTGGCCGCGGTGGTCTTGCTGGCCTTTGTCACCTGCTTCGCCCCCAAC
ArgAlaValGlyLeuAlaAlaValValLeuLeuAlaPheValThrCysPheAlaProAsn

           790                 810                 830
AACTTCGTGCTCCTGGCGCACATCGTGAGCCGCCTGTTCTACGGCAAGAGCTACTACCAC
AsnPheValLeuLeuAlaHisIleValSerArgLeuPheTyrGlyLysSerTyrTyrHis

           850                 870                 890
GTGTACAAGCTCACGCTGTGTCTCAGCTGCCTCAACAACTGTCTGGACCCGTTTGTTTAT
ValTyrLysLeuThrLeuCysLeuSerCysLeuAsnAsnCysLeuAspProPheValTyr

           910                 930                 950
TACTTTGCGTCCCGGGAATTCCAGCTGCGCCTGCGGGAATATTTGGGCTGCCGCCGGGTG
TyrPheAlaSerArgGluPheGlnLeuArgLeuArgGluTyrLeuGlyCysArgArgVal

           970                 990                1010
CCCAGAGACACCCTGGACACGCGCCGCGAGAGCCTCTTCTCCGCCAGGACCACGTCCGTG
ProArgAspThrLeuAspThrArgArgGluSerLeuPheSerAlaArgThrThrSerVal

          1030                1050                1070
CGCTCCGAGGCCGGTGCGCACCCTGAAGGGATGGAGGGAGCCACCAGGCCCGGCCTCCAG
ArgSerGluAlaGlyAlaHisProGluGlyMetGluGlyAlaThrArgProGlyLeuGln

          1090                1110                1130
AGGCAGGAGAGTGTGTTCTGAGTCCCGGGGGGCGCAGCTTGGAGAGCCGGGGGGCGCAGCTT
ArgGlnGluSerValPhe

          1150                1170                1190
GGAGATCCAGGGGGCGCATGGAGAGGCCACGGTGCCAGAGGTTCAGGGAGAA
```

Fig.3B

5/6

FIG.4

6/6
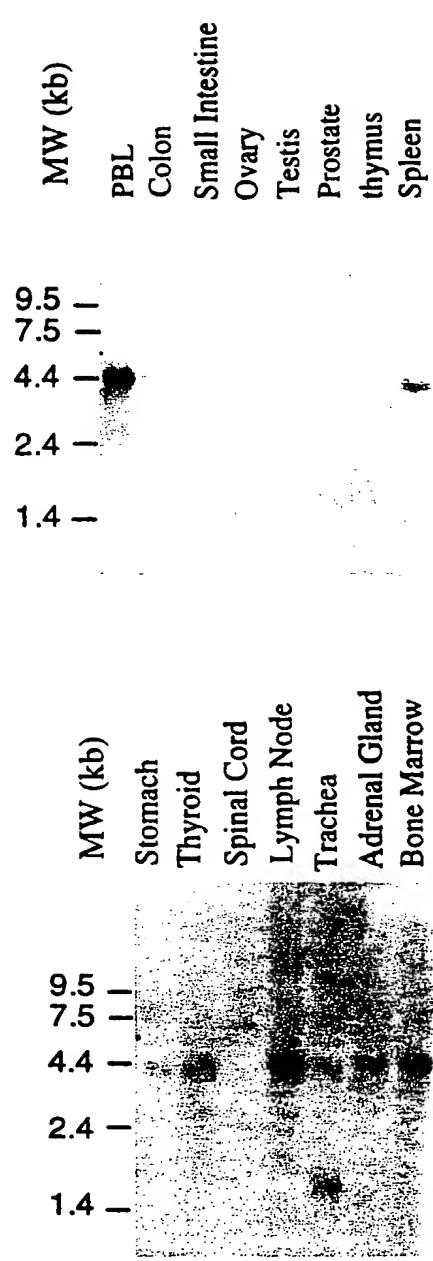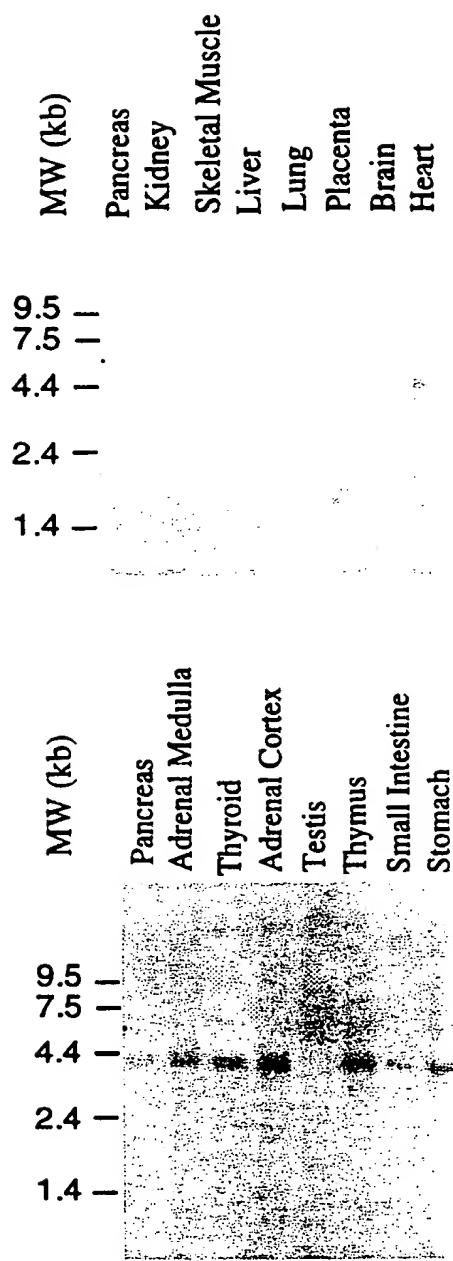
```
                                          10        20        30
HG52.pep                          MQVPNSTGPDNATLQMLRNPAIAVALPVVYS
                                  |:|   : |  :   :| :  ::: :| ||:
hTHRR    EPFWEDEEKNESGLTEYRLVSINKSSPLQKQLPAFISEDASG--YLTSSWLTLFVPSVYT
           60        70        80        90       100       110


                40        50        60        70        80        90
HG52.pep LVAANVIPGNLFSLWVLCRRMGPRSPSVIFMINLSVTDLMLASVLPFQIYYHCNRHHWVF
          | :||:| |:::: |:  :|   ::|:|::|::|:::|::::||||:|  |: :   | |
hTHRR    GVFVVSLPLNIMAIVVFILKMKVKKPAVVYMLHLATADVLFVSVLPFKISYYFSGSDWQF
           120         130       140       150       160       170


               100       110       120       130       140       150
HG52.pep GVLLCNVVTVAFYANMYSSILTMTCISVERFLGVLYPLSSKRWRRRRYAVAACAGTWLLL
         |   ||   ||:|||  |||:|||  ||  ||:::|||:|:||::|   ||     |   :|  :  | |
hTHRR    GSELCRFVTAAFYCNMYASILLMTVISIDRFLAVVYPMQSLSWRTLGRASFTCLAIWALA
           180       190       200       210       220       230


               160       170       180       190       200       210
HG52.pep LTALSPLARTDLTYPVHALGIITCFDVLKWTMLPSVAMWAVFLFTIF-ILLFLIPFVITV
         ::::  ||:  : |  | |:|:| || |||: |:| :    : :: |: |   ::|::|::|::
hTHRR    IAGVVPLVLKEQTIQVPGLNITTCHDVLNETLLEG---YYAYYFSAFSAVFFFVPLIIST
           240       250       260       270       280


               220       230       240       250       260
HG52.pep ACYTATILKLLRTEEAHGREQRRRAVGLAAVVLLAFVTCFAPNNFVLLAH--IVSRLFYG
         :||: :|:: | :  : :| :: ||: |:|:|: |: ||:|:| :|:|| ::|:
hTHRR    VCYV-SIIRCLSSSAVANRSKKSRALFLSAAVFCIFIICFGPTNVLLIAHYSFLSHTSTT
         290         300       310       320       330       340


         270       280       290       300       310       320
HG52.pep KSYYHVYKLTLCLSCLNNCLDPFVYYFASREFQLRLREYLGCRRVPRDTLDTRRESLFSA
         ::  | :| |  :|:| :::|:||::||:|| | |  :    | |::
hTHRR    EAAYFAYLLCVCVSSISSCIDPLIYYYASSECQRYVYSILCCKESSDPSSYNSSGQLMAS
         350         360       370       380       390       400


         330       340       350       360
HG52.pep RTTSVRSEAGAHPEGMEGATRPGLQRQESVF

hTHRR    KMDTCSSNLNNSIYKKLLT
         410       420
```

FIG.5